LIST OF SEQUENCES

(1) GENERAL INFORMATION:

(i) APPLICANT:

T,0240

(A) NAME: CYANAMID IBERICA, S.A.

(B) ADDRESS: CRISTOBAL BORDIU, 35

(C) CITY: MADRID

(E COUNTRY: SPAIN

(F) POSTAL CODE: 28003

(G TELEPHONE: 34 1 663 91 21

(H) TELEFAX: 34 1 663 94 01

(ii) TITLE OF THE INVENTION:

VECTORS BASED ON RECOMBINANT DEFECTIVE VIRAL GENOMES AND THEIR USE IN THE FORMULATION OF VACCINES

(iii) NUMBER OF SEQUENCES: 24

(iv) FORM READABLE BY COMPUTER:

(A) MEDIUM: Diskette

(B) COMPUTER: IBM PC-compatible

(C) OPERATIVE SYSTEM: PC-DOS/MS-DOS

(D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION ON IDENTIFIED SEQUENCE No. [ID. SEQ. No.]: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 27 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 1:

GTGAGTGTAG CGTGGCTATA TCTCTTC 27

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 2:

CCGTTGTGGT GTCACATTAA C 21

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 3:

GCCTCTAGAG GAGCTTTGTG GTTCACTTAC AC 32

Pat. 28.5.97

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 4:

GCTCTAGAGC GTTTGAATCA ACCCCCAAAA GC 32

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 5:

GGAATTCCGG GACTATCCTA AGTGTG 26

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 6:

GGAATTCCAG CAATACTATT ATCAA 25

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 7:

TTGATAATAG TATTGCTGGC 20

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 8:

GGACTAGTAT CACTATCAAA AGG 23

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 9:

GATGGATGTT GTGGTGTGAG 20

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 10:

CGAGTTGGTG TCCGAAGACA AAATCT 26

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 11:

ATACGAGCAT CAATATCACC 20

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 12:

AGAGTTGCCA CAGACTGCAG 20

(2) INFORMATION ON IDENTIFIED SEQUENCE No.: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 13:

CAGCAGTTC AAAGTTACCC 20

(2) INFORMATION ON IDENTIFIED SEQUENCE No. 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 14:

CCATTGTTAA GCCAACAACC 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 15 :

ATCACACTTA GGATAGTCCC 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 16 :

GTCTAACAAT GTGCCAAGG 19

2) INFORMATION ON IDENTIFIED SEQUENCE No. 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 17:

GCCAGCAATA CTATTATCAA 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 18:

CACTGTGGCA CCCTTACCTG 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 19:

GTACACCCAC TATGTTGTCT 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 20 :

TTGCGAGTGA AAACAAATGT 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 21:

CTCACAATCA GACGCTGTAC 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 22:

GACACGTTGT CCCTGGTTGG 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: monocatenary

(D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: DNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 23:

ACATTTTAAA CAATCACTAG 20

2) INFORMATION ON IDENTIFIED SEQUENCE No. 24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9714 base pairs

(B) TYPE: nucleic acid

(C) NO. OF CHAINS: bicatenary

(D) TOPOLOGY: circular

(ii) TYPE OF MOLECULE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Virus of transmissible gastroenteritis of pigs (TGEV)

(B) STRAIN: THER-1

(C) INDIVIDUAL ISOLATE: DI-C

(xi) DESCRIPTION OF THE SEQUENCE: ID. SEQ. No.: 24:

```
HCTTTTAAAG TAAAGTGAGT GTAGCGTGGC TATATCTCTT CTTTTACTTT AACTAGCCTT   60
GTGCTAGATT TTGTCTTCGG ACACCAACTC GAACTAAACG AAATATTTGT CTTTCTATGA  120
AATCATAGAG GACAAGCGTT GATTATTTCC ATTCAGTTTG GCAATCACTC CTTGGAACGG  180
GGTTGAGCGA ACGGTGCAGT AGGGTTCCGT CCCTATTTCG TAAGTCGCCT AGTAGTAGCG  240
AGTGCGGTTC CGCCCGTACA ACGTTGGGTA GACCGGGTTC CGTCCTGTGA TCTCCCTCGC  300
CGGCCGCCAG GAGAATGAGT TCCAAACAAT TCAAGATCCT TGTTAATGAG GACTATCAAG  360
TCAACGTGCC TAGTCTTCCT ATTCGTGACG TGTTACAGGA AATTAAGTAC TGCTACCGTA  420
ATGGATTTGA GGGCTATGTT TTCGTACCAG AATACTGTCG TGACCTAGTT GATTGCGATC  480
GTAAGGATCA CTACGTCATT GGTGTTCTTG GTAACGGAGT AAGTGATCTT AAACCTGTTC  540
TTCTTACCGA ACCCTCCGTC ATGTTGCAAG GCTTTATTGT TAGAGCTAAC TGCAATGGCG  600
TTCTTGAGGA CTTTGACCTT AAAATTGCTC GCACTGTCAG AGGTGCCATA TATGTTGATC  660
AATACATGTG TGGTGCTGAT GGAAAACCAG TCATTGAAGG CGATTTTAAG GACTACTTCG  720
GTGATGAAGA CATCATTGAA TTTGAAGGAG AGGAGTACCA TTGCGCTTGG ACAACTGTGC  780
GCGATGAGAA ACCGCTGAAT CAGCAAACTC TCTTTACCAT TCAGGAAATC CAATACAATC  840
TGGACATTCC TCATAAATTG CCAAACTGTG CTACTAGACA TGTAGCACCA CCAGTCAAAA  900
AGAACTCTAA AATAGTTCTG TCTGAAGATT ACAAGAAGCT TTATGATATC TTCGGATCAC  960
CCTTTATGGG AAATGGTGAC TGTCTTAGCA AATGCTTTGA CACTCTTCAT TTTATCGCTG 1020
CTACTCTTAG ATGCCCGTGT GGTTCTGAAA GTAGCGGCGT TGGAGATTGG ACTGGTTTTA 1080
AGACTGCCTG TTGTGGTCTT TCTGGCAAAG TTAAGGGTGT CACTTTGGGT GATATTAAGC 1140
CTGGTGATGC TGTTGTCACT AGTATGAGCG CAGGTAAGGG AGTTAAGTTC TTTGCCAATT 1200
GTGTTCTTCA ATATGCTGGT GATGTTGAAG GTGTCTCCAT CTGGAAAGTT ATTAAAACTT 1260
TTACAGTTGA TGAGACTGTA TGCACCCCTG GTTTTGAAGG CGAATTGAAC GACTTCATCA 1320
AACCTGAGAG CAAATCACTA GTTGCATGCA GCGTTAAAAG AGCATTCATT ACTGGTGATA 1380
TTGATGATGC TGTACATGAT TGTATCATTA CAGGAAAATT GGATCTTAGT ACCAACCTTT 1440
TTGGTAATGT TGGTCTATTA TTCAAGAAGA CTCCATGGTT TGTACAAAAG TGTGGTGCAC 1500
TTTTTGTAGA CGCTTGGAAA GTAGTAGAGG AGCTTTGTGG TTCACTCACA CTTACATACA 1560
AGCAAATTTA TGAAGTTGTA GCATCACTTT GCACTTCTGC TTTTACGATT GTAAACTACA 1620
AGCCAACATT TGTGGTTCCA GACAATCGTG TTAAAGATCT TGTAGACAAG TGTGTGAAAG 1680
TTCTTGTAAA AGCATTTGAT GTTTTTACGC AGATTATCAC AATAGCTGGT ATTGAGGCCA 1740
AATGCTTTGT GCTTGGTGCT AAATACCTGT TGTTCAATAA TGCACTTGTC AAACTTGTCA 1800
GTGTTAAAAT CCTTGGCAAG AAGCAAAAGG GTCTTGAATG TGCATTCTTT GCTACTAGCT 1860
TGGTTGGTGC AACTGTTAAT GTGACACCTA AAAGAACAGA GACTGCCACT ATCAGCTTGA 1920
ACAAGGTTGA TGATGTTGTA GCACCAGGAG AGGGTTATAT CGTCATTGTT GGTGATATGG 1980
CTTTCTACAA GAGTGGTGAA TATTATTTCA TGATGTCTAG TCCTAATTTT GTTCTTACTA 2040
ACAATGTTTT TAAAGCAGTT AAAGTTCCAT CTTATGACAT CGTTTATGAT GTTGATAATG 2100
```

ATACCAAAAG CAAAATGATT GCAAAACTTG GTTCATCATT TGAACAAATA CCAACTGGCA 2160
CACAAGATCC AATTCGGTTC TGTATTGAAA ATGAAGTTTG TGTTGTCTGT GGTTGTTGGC 2220
TTAACAATGG TTGCATGTGC GATCGTACTT CTATGCAGAG TTTTACTGTT GATCAAAGTT 2280
ATTTAAACGA GTGCGGGGTT CTAGTGCAGC TCGACTAGAA CCCTGCAATG GTACTGATCC 2340
AGACCATGTT AGTAGAGCTT TTGACATCTA CAACAAAGAT GTTGCGTGTA TTGGTAAATT 2400
CCTTAAGACG AATTGTTCAA GATTTAGGAA TTTGGACAAA CATGATGCCT ACTACATTGT 2460
CAAACGTTGT ACAAAGACCG TTATGGACCA TGAGCAAGTC TGTTATAACG ATCTTAAAGA 2520
TTCTGGTGCT GTTGCTGAGC ATGACTTCTT CACATATAAA GAGGGTAGAT GTGAGTTCGG 2580
TAATGTTGCA CGTAGGAATC TTACAAAGTA CACAATGATG GATCTTTGTT ACGCTATCAG 2640
AAATTTTGAT GAAAAGAACT GTGAAGTTCT CAAAGAAATA CTCGTGACAG TAGGTGCTTG 2700
CACTGAAGAA TTCTTTGAAA ATAAAGATTG GTTTGATCCA GTTGAAAATG AAGCCATACA 2760
TGAAGTTTAT GCAAAACTTG GACCCATTGT AGCCAATGCT ATGCTTAAAT GTGTTGCTTT 2820
TTGCGATGCG ATAGTGGAAA AAGGCTATAT AGGTGTTATA ACACTTGACA ACCAAGATCT 2880
TAATGGCAAT TTCTACGATT TCGGCGATTT CGTGAAGACT GCTCCGGGTT TTGGTTGCGC 2940
TTGTGTTACA TCATATTATT CTTATATGAT GCCTTTAATG GGGATGACTT CATGCTTAGA 3000
GTCTGAAAAC TTTGTGAAAA GTGACATCTA TGGTTCTGAT TATAAGCAGT ATGATTTACT 3060
AGCTTATGAT TTTACCGAAC ATAAGGAGTA CCTTTTCCAA AAATACTTTA AGTACTGGGA 3120
TCGCACATAT CACCCAAATT GTTCTGATTG TACTAGTGAG GAGTGTATTA TTCATTGTGC 3180
TAATTTTAAC ACATTGTTTT CTATGACAAT ACCAATGACA GCTTTTGGAC CACTTGTCCG 3240
TAAAGTTCAT ATTGATGGTG TACCAGTAGT TGTTACTGCA GGTTACCATT TCAAACAACT 3300
TGGTATAGTA TGGAATCTTG ATGTAAAATT AGACACAATG AAGTTGAGCA TGACTGATCT 3360
TCTTAGATTT GTCACAGATC CAACACTTCT TGTAGCATCA AGCCCTGCAC TTTTAGACCA 3420
GCGTACTGTC TGTTTCTCCA TTGCAGCTTT GAGTACTGGT ATTACATATC AGACAGTAAA 3480
ACCAGGTCAC TTTAACAAGG ATTTCTACGA TTTCATAACA GAGCGTGGAT TCTTTGAAGA 3540
GGGATCTGAG TTAACATTAA AACATTTTTT CTTTGCACAG GGTGGTGAAG CTGCTATGAC 3600
AGACTTCAAT TATTATCGCT ACAATAGAGT CACAGTACTT GATATTTGCC AAGCTCAATT 3660
TGTTTACAAA ATAGTTGGCA AGTATTTTGA ATGTTATGAC GGTGGGTGCA TTAATGCTCG 3720
TGAAGTTGTT GTTACAAACT ATGACAAGAG TGCTGGCTAT CCTTTGAACA AATTTGGTAA 3780
AGCTAGACTT TACTACGAAA CTCTTTCATA TGAAGAGCAG GATGCACTTT TTGCTTTAAC 3840
AAAGAGAAAT GTTTTACCCA CAATGACTCA AATGAATTTG AAATACGCTA TTTCTGGTAA 3900
GGCAAGAGCT CGTACAGTAG GAGGAGTTTC ACTTCTTTCT ACCATGACTA CGAGACAATA 3960
TCATCAGAAG CATTTGAAGT CAATTGCTGC AACACGCAAT GCTACTGTGG TCATTGGTTC 4020
AACCAAGTTT TATGGTGGTT GGGACAATAT GCTTAAAAAT TTAATGCGTG ATGTTGATAA 4080
TGGTTGTTTG ATGGGATGGG ACTATCCTAA GTGTGACCGT GCTTTACCTA ATATGATTAG 4140
AATGGCTTCT GCCATGATAT TAGGTTCTAA GCATGTTGGT TGTTGTACAC ATAATGATAG 4200

GTTCTACCGC CTCTCCAATG AGTTAGCTCA AGTACTCACA GAAGTTGTGC ATTGCACAGG 4260

TGGTTTTTAT TTTAAACCTG GTGGTACAAC TAGCGGTGAT GGTACTACAG CATATGCTAA 4320

CTCTGCTTTT AACATCTTTC AAGCTGTTTC TGCTAATGTT AATAAGCTTT TGGGGGTTGA 4380

TTCAAACGCT TGTAACAACG TTACAGTAAA ATCCATACAA CGTAAAATTT ACGATAATTG 4440

TTATCGTAGT AGCAGCATTG ATGAAGAATT TGTTGTTGAG TACTTTAGTT ATTTGAGAAA 4500

ACACTTTTCT ATGATGATTT TATCTGATGA TGGAGTTGTG TGCTACAACA AAGATTATGC 4560

GGATTTAGGT TATGTAGCTG ACATTAATGC TTTTAAAGCA ACACTTTATT ACCAGAATAA 4620

CGTCTTTATG TCCACTTCTA AGTGTTGGGT AGAACCAGAT CTTAGTGTTG GACCACATGA 4680

ATTTTGTTCA CAGCATACAT TGCAGATTGT TGGGCCTGAT GGAGACTACT ATCTTCCCTA 4740

TCCAGACCCG TCCAGAATTT TGTCAGCTGG TGTGTTTGTT GATGACATAG TTAAAACAGA 4800

CAATGTTATT ATGTTAGAAC GTTACGTGTC ATTGGCTATT GACGCATACC CGCTCACAAA 4860

ACACCCTAAG CCTGCTTATC AAAAAGTGTT TTACACTCTA CTAGATTGGG TTAAACATCT 4920

ACAGAAAAAT TTGAATGCAG GTGTTCTTGA TTCGTTTTCA GTGACAATGT TAGAGGAAGG 4980

TCAAGATAAG TTCTGGAGTG AAGAGTTTTA CGCTAGCCTC TATGAAAAGT CCACTGTCTT 5040

GCAAGCTGCA GGCATGTGTG TAGTATGTGG TTCGCAAACT GTACTTCGTT GTGGAGACTG 5100

TCTTAGGAGA CCACTTTTAT GCACGAAATG TGCTTACGAC CATGTTATGG GAACAAAGCA 5160

TAAATTCATT ATGTCTATCA CACCATATGT GTGTAGTTTT AATGGTTGTA ATGTCAATGA 5220

TGTTACAAAG TTGTTTTTAG GTGGTCTTAG TTATTATTGT ATGAACCACA AACCACAGTT 5280

GTCATTCCCA CTCTGTGCTA ATGGCAACGT TTTTGGTCTA TATAAAAGTA GTGCAGTCGG 5340

CTCAGAGGCT GTTGAAGATT TCAACAAACT TGCAGTTTCT GACTGGACTA ATGTAGAAGA 5400

TTACAAACTT GCTAACAATG TCAAGGAATC TCTGAAAATT TTCGCTGCTG AAACTGTGAA 5460

AGCTAAGGAG GAGTCTGTTA AATCTGAATA TGCTTATGCT GTATTAAAGG AGGTTATCGG 5520

CCCTAAGGAA ATTGTACTCC AATGGGAAGC TTCTAAGACT AAGCCTCCAC TTAACAGAAA 5580

TTCAGTTTTC ACGTGTTTTC AGATAAGTAA GGATACTAAA ATTCAATTAG GTGAATTTGT 5640

GTTTGAGCAA TCTGAGTACG GTAGTGATTC TGTTTATTAC AAGAGCACGA GTACTTACAA 5700

ATTGACACCA GGTATGATTT TTGTGTTGAC TTCTCATAAT GTGAGTCCTC TTAAAGCTCC 5760

AATTTTAGTC AACCAAGAAA AGTACAATAC CATATCTAAG CTCTATCCTG TCTTTAATAT 5820

AGCGGAGGCC TATAATACAC TGGTTCCTTA CTACCAAATG ATAGGTAAGC AAAAATTTAC 5880

AACTATCCAA GGTCCTCCTG GTAGCGGTAA ATCTCATTGT GTTATAGGTT TGGGTTTGTA 5940

TTACCCTCAG GCGAGAATAG TCTACACTGC ATGTTCTCAT GCGGCTGTAG ACGCTTTATG 6000

TGAAAAAGCA GCCAAAAACT TCAATGTTGA TAGATGTTCA AGGATAATAC CTCAAAGAAT 6060

CAGAGTTGAT TGTTACACAG GCTTTAAGCC TAATAACACC AATGCGCAGT ACTTGTTTTG 6120

TACTGTTAAT GCTCTACCAG AAGCAAGTTG TGACATTGTT GTAGTTGATG AGGTCTCTAT 6180

GTGTACTAAT TATGATCTTA GTGTCATAAA TAGCCGACTG AGTTACAAAC ATATTGTTTA 6240

TGTTGGAGAC CCACAGCAGC TACCAGCTCC TAGAACTTTG ATTAATAAGG GTGTACTTCA 6300

ACCGCAGGAT TACAATGTTG TAACCAAAAG AATGTGCACA CTAGGACCTG ATGTCTTTTT 6360
GCATAAATGT TACAGGTGCC CAGCTGAAAT TGTTAAAACA GTCTCTGCAC TTGTTTATGA 6420
AAATAAATTT GTACCTGTCA ACCCAGAATC AAAGCAGTGC TTCAAAATGT TTGTAAAAGG 6480
TCAGATTCAG ATTGAGTCTA ACTCTTCTAT AAACAACAAG CAACTAGAGG TTGTCAAGGC 6540
CTTTTTAGCA CATAATCCAA AATGGCGTAA AGCTGTTTTC ATCTCACCCT ATAATAGTCA 6600
AAATTATGTT GCTCGGCGTC TTCTTGGTTT GCAAACGCAA ACTGTGGATT CCGCTCAGGG 6660
TAGTGAGTAT GATTACGTCA TCTAGCTGCT CTGAAGATTT TTAATCCTGC TGCAATTCAC 6720
GATGTGGGTA ATCCAAAAGG CATCCGTTGT GCTACAACAC CAATACCATG GTTTTGTTAT 6780
GATCGTGATC CTATTAATAA CAATGTTAGA TGTCTGGATT ATGACTATAT GGTACATGGT 6840
CAAATGAATG GTCTTATGTT ATTTTGGAAC TGTAATGTAG ACATGTACCC AGAGTTTTCA 6900
ATTGTTTGTA GATTTGATAC TCGCACTCGC TCTAAATTGT CTTTAGAAGG TTGTAATGGT 6960
GGTGCATTGT ATGTTAATAA CCATGCTTTC CACACACCAG CTTATGATAG AAGAGCTTTT 7020
GCTAAGCTTA AACCTATGCC ATTCTTTTAC TATGATGATA GTAATTGTGA ACTTGTTGAT 7080
GGGCAACCTA ATTATGTACC ACTTAAGTCA AATGTTTGCA TAACAAAATG CAACATTGGT 7140
GGTGCTGTCT GCAAGAAGCA TGCTGCTCTT TACAGAGCGT ATGTTGAGGA TTACAACATT 7200
TTTATGCAGG CTGGTTTTAC AATATGGTGT CCTCAAAACT TTGACACCTA TATGCTTTGG 7260
CATGGTTTTG TTAATAGCAA AGCACTTCAG AGTCTAGAAA ATGTGGCTTT TAATGTCGTT 7320
AAGAAAGGTG CCTTCACCGG TTTAAAAGGT GACTTACCAA CTGCTGTTAT TGCTGACAAA 7380
ATAATGGTAA GAGATGGACC TACTGACAAA TGTATTTTTA CAAATAAGAC TAGTTTACCT 7440
ACAAATGTAG CTTTTGAGTT ATATGCAAAA CGCAAACTTG GACTCACACC TCCATTAACA 7500
ATACTTAGGA ATTTAGGTGT TGTCGCAACA TATAAGTTTG TGTTGTGGGA TTATGAAGCT 7560
GAACGTCCTT TCTCAAATTT CACTAAGCAA GTGTGTTCCT ACACTGATCT TGATAGTGAA 7620
GTTGTAACAT GTTTTGATAA TAGTATTGCT GGCTCTTTTG AGCGTTTTAC TACTACAAGA 7680
GATGCAGTGC TTATTTCTAA TAACGCTGTG AAAGGGCTTA GTGCCATTAA ATTACAATAT 7740
GGCCTTTTGA ATGATCTACC TGTAAGTACT GTTGGAAATA AACCTGTCAC ATGGTATATC 7800
TATGTGCGCA AGAATGGTGA GTACGTCGAA CAAATCGATA GTTACTATAC ACAGGGACGT 7860
ACTTTTGAAA CCTTCAAACC TCGTAGTACA ATGGAAGAAG ATTTTCTTAG TATGGATACT 7920
ACACTCTTCA TCCAAAAGTA TGGTCTTGAG GATTATGGTT TTGAACACGT TGTATTTGGA 7980
GATGTCTCTA AAACTACCAT TGGTGGTATG CATCTTCTTA TATCGCAAGT GCGCCTTGCA 8040
AAAATGGGTT TGTTTTCCGT TCAAGAATTT ATGAATAATT CTGACAGTAC ACTGAAAAGT 8100
TGTTGTATTA CATATGCTGA TGATCCATCT TCTAAGAATG TGTGCACTTA TATGGACATA 8160
CTCTTGGACG ATTTTGTGAC TATCATTAAG AGCTTAGATC TTAATGTTGT GTCCAAAGTT 8220
GTGGATGTCA TTGTAGATTG TAAGGCATGG AGATGGATGT TGTGGTGTGA GAATTCACAT 8280
ATTAAAACCT TCTATCCACA ACTCCAATCT GCTGAATGGA ATCCCGGCTA AGCATGCCT 8340
ACACTGTACA AAATCCAGCG TATGTGTCTC GAACGGTGTA ATCTCTACAA TTATGGTGCA 8400

CAAGTGAAAT TACCTGATGG CATTACTACT AATGTCGTTA AGTATACTCA GTTGTGTCAA 8460
TACCTTAACA CTACTACATT GTGTGTACCA CACAAAATGC GTGTATTGCA TTTAGGAGCT 8520
GCTGGTGCAT CTGGTGTTGC TCCTGGTAGT ACTGTATTAA GAAGATGGTT ACCAGATGAT 8580
GCCATATTGG TTGATAATGA TTTGAGAGAT TACGTTTCCG ACGCAGACTT CAGTGTTACA 8640
GGTGATTGTA CTAGTCTTTA CATCGAAGAC AAGTTTGATT TGCTCGTCTC TGATTTATAT 8700
GATGGCTCCA CAAAATCAAT TGACGGTGAA AACACGTCGA AAGATGGTTT CTTTACTTAT 8760
ATTAATGGTT TCATTAAAGA GAAACTGTCA CTTGGTGGAT CTGTTGCCAT TAAAATCACG 8820
GAATTTAGTT GGAATAAAGA TTTATATGAA TTGATTCAAA GATTTGAGTA TTGGACTGTG 8880
TTTTGTACAA GTGTTAACAC GTCATCATCA GAAGGCTTTC TGATTGGTAT TAACTACTTA 8940
GGACCATACT GTGACAAAGC AATAGTAGAT GGAAATATAA TGCATGCCAA TTATATATTT 9000
TGGAGAAACT CTACAATTAT GGCTCTATCA CATAACTCAG TCCTAGACAC TCCTAAATTC 9060
AAGTGTCGTT GTAACAACGC ACTTATTGTT AATTTAAAAG AAAAAGAATT GAATGAAATG 9120
GTCATTGGAT TACTAAGGAA GGGTAAGTTG CTCATTAGAA ATAATGGTAA GTTACTAAAC 9180
TTTGGTAACC ACTTCGTTAA CACACCATGA AAAAATGCTG TATTTATTAC AGTTTTAATC 9240
TTACTACTAA TTGGTAGACT CCAATTATTA GAAAGACTAT TACTTAATCA CTCTTTCAAT 9300
CTTAAAACTG TCAATGACTT TAATATCTTA TATAGGAGTT TAGCAGAAAC CAGATTACTA 9360
AAAGTGGTGC TTCGAGTAAT CTTTCTAGTC TTACTAGGAT TTTGCTGCTA CAGATTGTTA 9420
GTCACATTAA TGTAAGGCAA CCCGATGTCT AAAACTGGTT TTTCCGAGGA ATTACTGGTC 9480
ATCGCGCTGT CTACTCTTGT ACAGAATGGT AAGCACGTGT AATAGGAGGT ACAAGCAACC 9540
CTATTGCATA TTAGGAAGTT TAGATTTGAT TTGGCAATGC TAGATTTAGT AATTTAGAGA 9600
AGTTTAAAGA TCCGCTACGA CGAGCCAACA ATGGAAGAGC TAACGTCTGG ATCTAGTGAT 9660
TGTTTAAAAT GTAAAATTGT TTGAAAATTT TCCTTTTGAT AGTGATACAA AAAA 9714